



**ARTICLE**

# Mapping of Land Use and Land Cover (LULC) Using EuroSAT and Transfer Learning

Suman Kunwar<sup>1,\*</sup> and Jannatul Ferdush<sup>2</sup>

<sup>1</sup>Faculty of Computer Science, Selinus University of Sciences and Literature, Ragusa, Italy

<sup>2</sup>Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore, Bangladesh

\*Corresponding Author: Suman Kunwar. Email: sumn2u@gmail.com

Received: 12 November 2023 Accepted: 28 December 2023 Accepted: 27 February 2024

## ABSTRACT

As the global population continues to expand, the demand for natural resources increases. Unfortunately, human activities account for 23% of greenhouse gas emissions. On a positive note, remote sensing technologies have emerged as a valuable tool in managing our environment. These technologies allow us to monitor land use, plan urban areas, and drive advancements in areas such as agriculture, climate change mitigation, disaster recovery, and environmental monitoring. Recent advances in Artificial Intelligence (AI), computer vision, and earth observation data have enabled unprecedented accuracy in land use mapping. By using transfer learning and fine-tuning with red-green-blue (RGB) bands, we achieved an impressive 99.19% accuracy in land use analysis. Such findings can be used to inform conservation and urban planning policies.

## KEYWORDS

Land use land cover; EuroSAT; transfer learning; convolutional neural networks; vision transformers

## 1 Introduction

The world population has increased significantly over the last few centuries and is projected to continue growing [1]. With growing demand, human beings are consuming more natural resources, including water, energy, minerals, and agricultural products. Activities such as agriculture, forestry, and urbanization contribute to 23% of global greenhouse gas emissions [2], primarily due to deforestation and land degradation. Monitoring land use changes is vital for better environmental management, urban planning, and nature protection [3,4]. Advancements in remote sensing have democratized access to satellite imagery, making it an open resource for anyone, fueling innovation and entrepreneurship. Thanks to the use of low orbit and geostationary satellites [5], we can observe the Earth with unprecedented detail. Moreover, improvements in remote sensing technology have resulted in better spatial resolution [6], enabling more precise ground surface analyses.

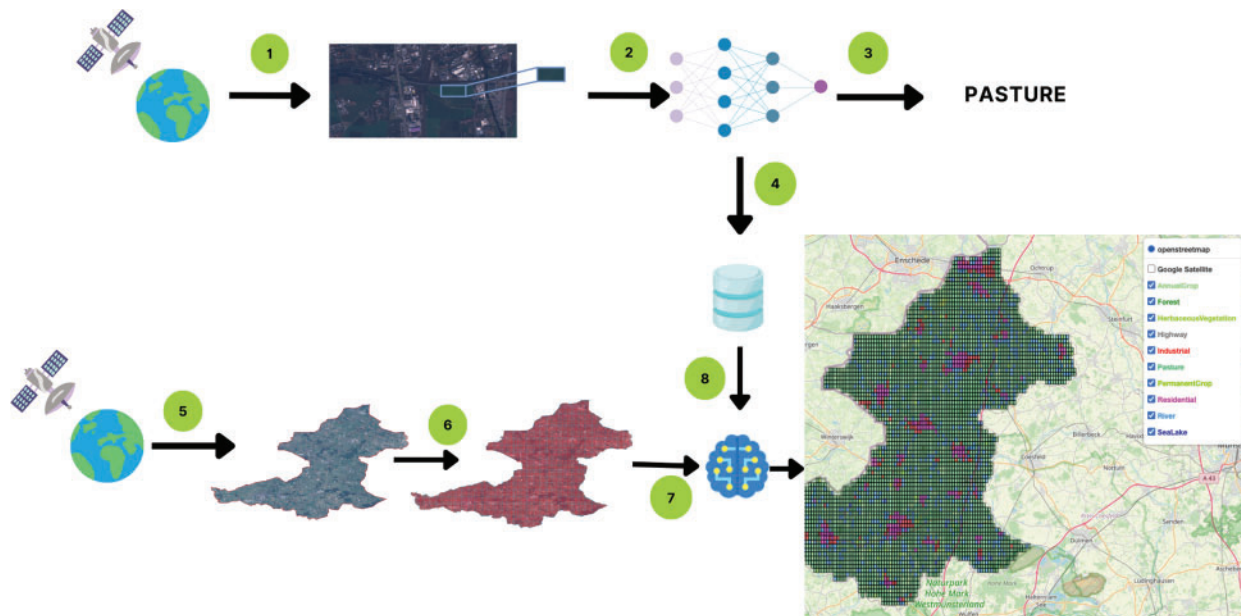
Such data access has fueled advancements in agriculture, urban development, climate change mitigation, disaster recovery, and environmental monitoring [7–9]. Advances in computer vision, AI, and earth observation data facilitate large scale land use mapping [10,11]. The community has



extensively embraced methods for classifying land use and land cover (LULC), including machine learning [12] and deep learning (DL) [13].

Recent studies suggest that DL techniques demonstrate remarkable performance in remote sensing (RS) image scene classification [14]. The objective of image scene classification and retrieval is to automatically allocate class labels to every RS image scene stored in an archive. This differs from semantic segmentation tasks used for LULC mapping and classification. One application of this technology is scene classification [15], which involves labeling an image based on specific semantic categories. This approach has numerous practical uses, including LULC analysis as well as land resource management [16,17].

Fig. 1 provides an overview of the LULC classification process using satellite images. Satellites capture images of the Earth. These images are then utilized to extract patches for classification. The objective is to automatically label the physical nature of land or its utilization. Each image patch is fed into a classifier, which then outputs the corresponding class depicted on the patch.



**Figure 1:** Mapping of LULC using satellite imagery

However, DL models tend to overfit and demand enormous quantities of labeled input data to perform well on unseen data [18]. This limitation has restricted their adoption in geosciences and remote sensing. Leveraging pre-trained models from optical datasets, such as ImageNet [19], can facilitate training new RS models using smaller labeled datasets. Various approaches have leveraged pretrained models alongside EuroSAT and have demonstrated encouraging results [20–22].

With the advancement of Vision Transformers (ViT), many applications are adopting it for image classification tasks [23], including EuroSAT [24,25]. It is suggested that further scaling can enhance performance [26], but this model has yet to be integrated with Geospatial data. The contributions of this paper are summarized below:

- We conduct a comprehensive evaluation of the ViT model on EuroSAT RGB data, with diverse settings and hyperparameters.
- We compare the effectiveness of ViT models trained using both augmented and non-augmented data. Additionally, we benchmark the performance of ViT against established models such as ResNet-50 and VGG16, providing valuable insights into its competitiveness in the field of LULC classification.
- We demonstrate the practical application of the best-performing ViT model by mapping LULC in the 'Kreis Borken' area using geospatial data. This visualization showcases the potential of ViT for real-world LULC analysis and monitoring.
- Our study's results add value to the continuous research on LULC classification through the application of deep learning models. The insights gained from this study can guide researchers and practitioners in developing and implementing more accurate and efficient LULC mapping techniques.

The paper is structured as follows: [Section 2](#) reviews related work. [Section 3](#) describes the dataset used in this study and introduces the methodologies applied using the ViT model. [Section 4](#) presents the results and provides an analysis. [Section 5](#) discusses the findings and concludes the paper in [Section 6](#).

## 2 Related Works

This section discusses various state-of-the-art image classification techniques that use DL and Transfer Learning (TL) for LULC using the EuroSAT dataset. Fine-tuning large-scale pretrained ViTs has shown prominent performance for computer vision tasks, such as image classification by Dosovitskiy et al. [27] using EuroSAT data from its early proposal. There is a growing need to adapt to different domains due to a shortage of data [28]. Several methods have been developed to tackle this challenge, including adversarial learning [29], feature representation learning [30], and unsupervised domain adaptation (UDA) [31]. To deal with the problem of limited labeled data for specific classes, new techniques such as few-shot learning (FSL) [32] and zero-shot learning (ZSL) [33] have been introduced. FSL enables models to generalize to new classes with just a few labeled examples, whereas ZSL can classify previously unseen classes without labeled examples by utilizing side information. These methods make use of prior knowledge from previous classes to learn about new ones, making them promising solutions for restricted data in LULC classification. A comparison of various spectral bands—RGB, RGB combined with Near InfraRed (NIR), and multispectral satellite images using the ViT model with Sentinel-2 EuroSAT was conducted by Anil et al. [34]. Their experimental results indicate that combining the NIR band with RGB provides more accurate results [35]. Various Deep Learning models including GoogleNet, ResNet-50, Random Forest, VGG, and CNN have been utilized for LULC classification.

In their study, Helber et al. [36] adopted a comprehensive approach, evaluating the performance of GoogleNet and ResNet-50 on various band combinations within the context of LULC classification. Among these, the ResNet-50 model demonstrated significant potential when trained on RGB bands, outperforming all other configurations investigated. Li et al.'s DDRL-AM method [37] achieved a peak accuracy of 98.74% with RGB bands. Yassine et al. [38] implemented two approaches: first, using the raw 13 Sentinel-2 bands on the EuroSAT dataset, reaching a respectable 98.78% accuracy. Then, they took it up a notch by adding calculated indices, propelling the accuracy to a commanding 99.58%, compared to others listed in [Table 1](#).

**Table 1:** Comparative analysis of studies for LULC classification with EuroSAT dataset

Authors	Model	Bands	Accuracy
Helber et al. [36]	GoogleNet	RGB	98.18%
Helber et al. [36]	ResNet-50	RGB	98.57%
Li et al. [37]	DDRL-AM	RGB	98.74%
Yassine et al. [38]	CNN	All 13 spectral bands	98.78%
Yassine et al. [38]	CNN	All 13 spectral bands + VIRE + NNIR + BR	99.58%
Naushad et al. [39]	WRN	RGB	99.17%

Naushad et al. [39] achieved a remarkable 99.17% accuracy by combining transfer learning with pre-trained VGG16 and Wide Residual Networks (WRNs) models on the RGB EuroSAT data. Techniques such as data augmentation [40], gradient clipping [41], adaptive learning rates [42], and early stopping [43] optimized performance and reduced computational time. Their WRN-based approach achieved a remarkable accuracy of 99.17%, setting a new benchmark in efficiency and accuracy. Recently, other variants, including hierarchical ViTs with diverse resolutions and spatial embeddings [44], have been proposed. Without a doubt, the advancements in large ViTs underscore the importance of developing efficient model adaptation strategies.

### 3 Materials and Methods

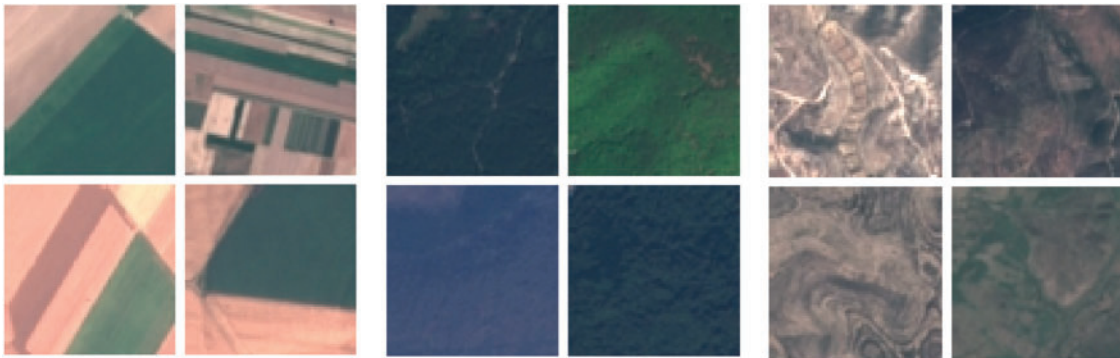
To classify LULC in a specific region using geospatial data, a TL task was undertaken. The EuroSAT classes were subsequently color-mapped with the assistance of the ViT model pre-trained on ImageNet-21k. Two datasets were utilized—one with data augmentation and another without. Taking advantage of the rich visual information in the RGB channels of the EuroSAT dataset, we trained our model using the PyTorch framework. Both model training and testing were conducted using the Tesla T4 Graphic Processing Units (GPUs) available on Google Colab.

#### 3.1 Dataset

EuroSAT dataset is considered novel and comprises 27,000 labeled and georeferenced images taken from the Sentinel-2 satellite. The images are classified into ten scene classes: Forest, Herbaceous Vegetation, Highway, Pasture, River, Industrial, Permanent Crop, Residential, Annual Crop, and Sea/Lake. Each image patch consists of  $64 \times 64$  pixels, featuring a spatial resolution of 10 meters. Fig. 2 shows some of the images from the EuroSAT dataset [45]. To train and evaluate our model effectively, we divided the dataset into two parts: a training set encompassing 80% of the data and a test set comprising the remaining 20%, both chosen randomly.

#### 3.2 Training and Evaluation

During model training, two sets of data were fed into the system: one with augmentation and other without. In the augmentation process, various image transformation techniques [46], such as crops, horizontal flips, and vertical flips, were utilized to augment the data. This strategy aids in preventing the neural network from overfitting to the training dataset, enabling it to generalize more effectively to unseen test data. Since we utilized pre-trained models, the input dataset was normalized to match the statistics (mean and standard deviation) of those models.



**Figure 2:** Sample images of three different classes from the EuroSAT dataset

The models were trained under the same settings, and their accuracy was subsequently evaluated using classification accuracy, which is calculated as the ratio of correct predictions to the total number of input samples [47]. F1-score is also used as an evaluation metric to measure the model accuracy by combining precision and recall scores. To reduce the computational cost and generalization we have limited the training to 15 epochs. The model with the highest accuracy was then tested with various epochs and its accuracy was measured. In this process, the loss was measured using cross-entropy loss. To counteract potential issues of vanishing or exploding gradients during training, which could adversely affect the parameters, the gradient clipping technique [48] was employed with a value set to 1.0. We used the Adam optimizer, combined with multiple learning rates, due to its proven efficacy in image classification tasks [49]. Regularization strategies, including early stopping, dropout, and weight decay [50], were also implemented to combat overfitting and optimize time and resources. We also recorded the total duration of the experiment.

## 4 Results

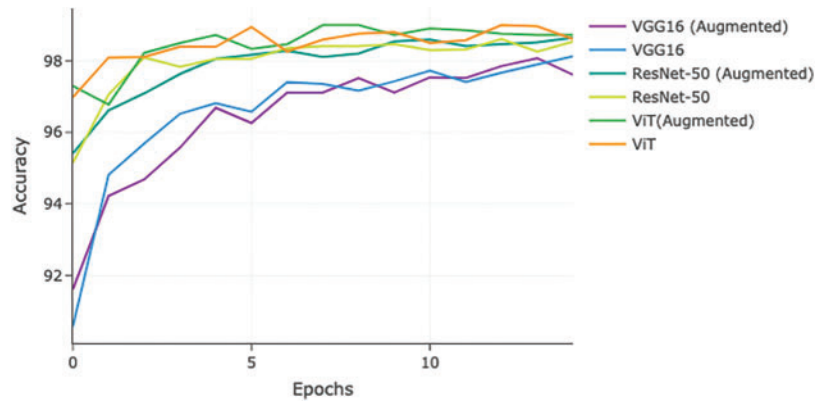
This section presents the results obtained from ViT, VGG16, and ResNet-50 models with same settings and narrows down the ViT model performance under different settings. Metrics such as accuracy and the time taken to train the model are measured. The model's performance is assessed using both data augmentation techniques and without data augmentation, maintaining similar settings for both.

Table 2 shows that the ViT model is more accurate in augmented and non-augmented data, but takes longer to train than other models. Conversely, the ResNet-50 model exhibits better accuracy than VGG16, and it requires less training time for both augmented and non-augmented data. The F1-score of ViT and ResNet are 0.99 whereas the VGG16 has 0.98 value in both augmented and non-augmented data. The loss and accuracy of both the augmented and non-augmented data of different models are depicted in Figs. 3 and 4.

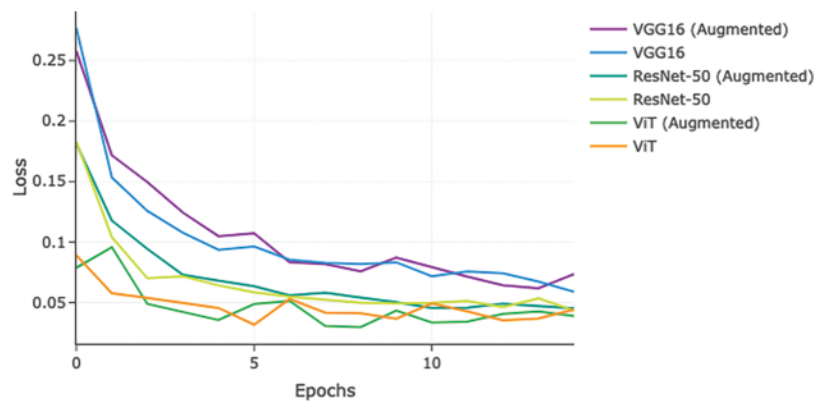
Based on the data presented in Table 3, it is clear that the accuracy of the ViT model improves significantly when augmented data is used and training time is increased. However, the accuracy declines for the non-augmented version. The F1-score also improves as training duration increases, but it drops suddenly for non-augmented data after 20 epochs. It is worth mentioning that training with augmented data takes longer than non-augmented data.

**Table 2:** Comparative experimental results of ViT, ResNet-50 and VGG16 models with and without augmented data

Model	Epochs	Total time (s)	F1-score	Accuracy
ViT	15	12427.12	0.99	98.61%
ViT (Augmented data)	15	11045.71	0.99	98.72%
VGG16	15	5070.30	0.98	98.06%
VGG16 (Augmented data)	15	4996.54	0.98	97.78%
ResNet-50	15	3602.74	0.99	98.52%
ResNet-50 (Augmented data)	15	3329.52	0.99	98.48%



**Figure 3:** Comparison of accuracy of different models

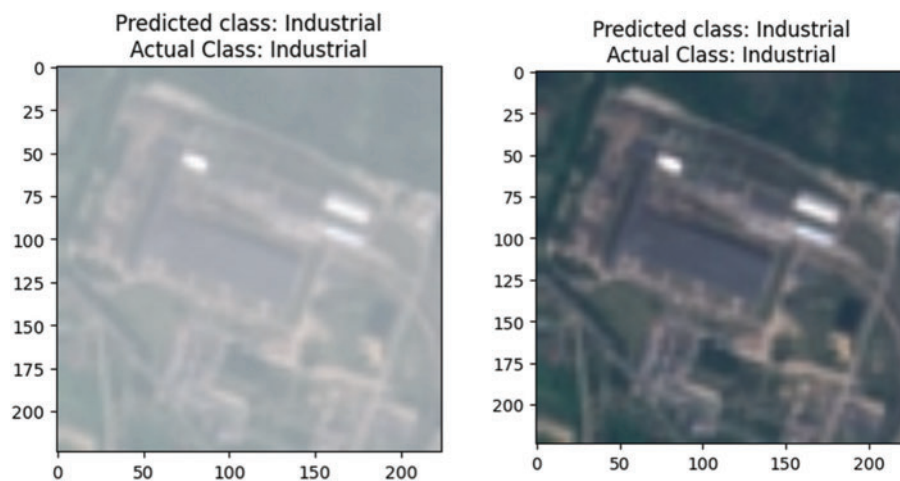


**Figure 4:** Comparison of loss of different models

**Table 3:** Comparative experimental results of ViT model with and without augmented data

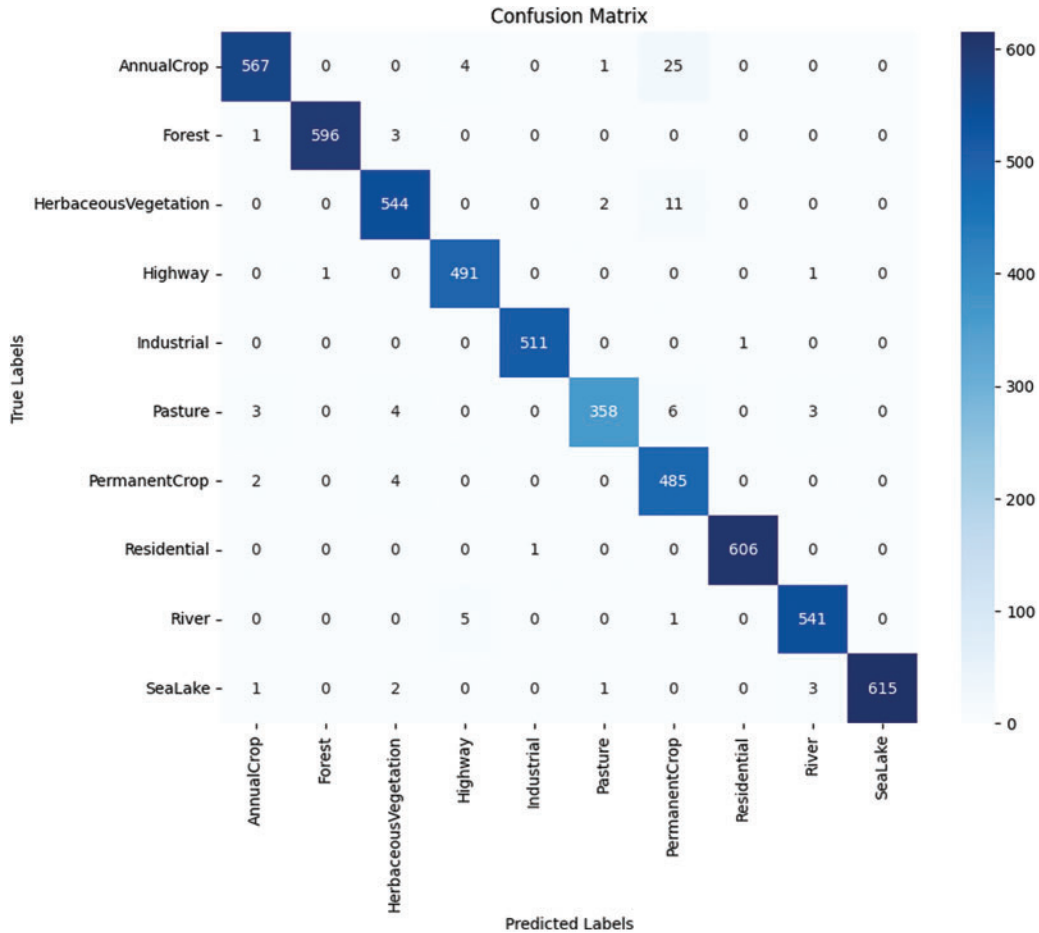
Model	Epochs	Total time (s)	F1-score	Accuracy
ViT	10	11045.71	0.98	98.67%
ViT (Augmented data)	10	7356.24	0.98	98.24%
ViT	15	12427.12	0.99	98.61%
ViT (Augmented data)	15	11045.71	0.99	98.72%
ViT	20	15604.66	0.98	98.41%
ViT (Augmented data)	20	14377.87	0.99	99.19%

Subsequently, the model is tested with test data, as shown in Fig. 5. Figs. 6 and 7 show the confusion matrix of ViT model with and without data augmentation on validation data. Analyzing the results, we found that data augmentation significantly improved the model's performance on certain classes, especially Forest and Sea Lake. The Pasture classes have the least accuracy, while the rest have almost 99% accuracy.



**Figure 5:** Correct prediction of test results for both augmented (left) and non-augmented (right) model using the ViT model trained at 20 epochs

From the experiments conducted, we found that the model trained with augmented data yields better accuracy. Using Google Earth Engine and Sentinel 2A [51], we selected the Kreis Borken area. The data, drawn from satellite images spanning 2018 to 2020, was segmented into  $64 \times 64$  tiles within the specified boundary. Using the model, these tiles were classified and color-mapped, as depicted in Fig. 8.



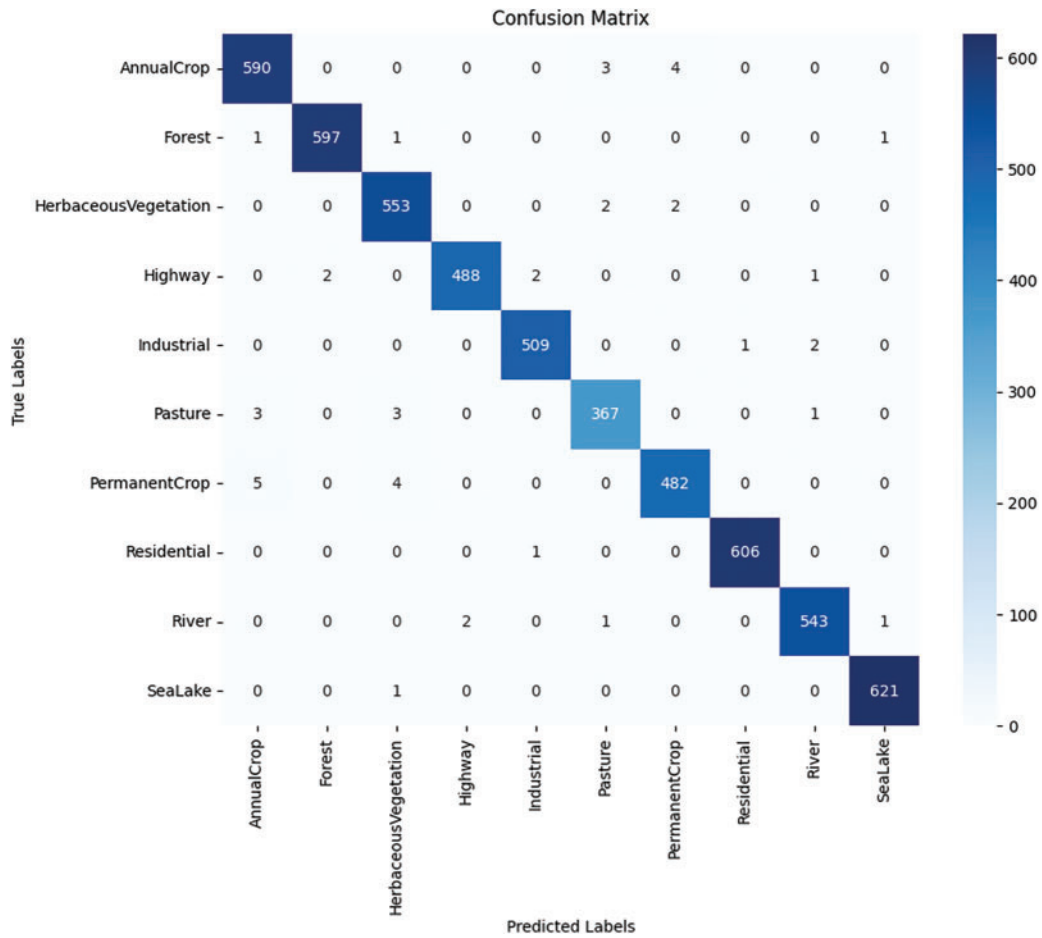
**Figure 6:** Confusion matrix of augmented data with 20 epochs of ViT model

## 5 Discussion

Our study aimed to investigate the effectiveness of transfer learning in LULC classification using a ViT model pre-trained on ImageNet-21K. We compared the performance of different models, including ViT, VGG16, and ResNet-50, and evaluated their classification accuracy, F1-score, and the time taken to train the models using both augmented and non-augmented datasets. We found that the Vision Transformer (ViT) model has higher accuracy and training time than the VGG16 and ResNet-50 models. However, the ResNet-50 model exhibits better accuracy than VGG16, and it requires less training time for both augmented and non-augmented data.

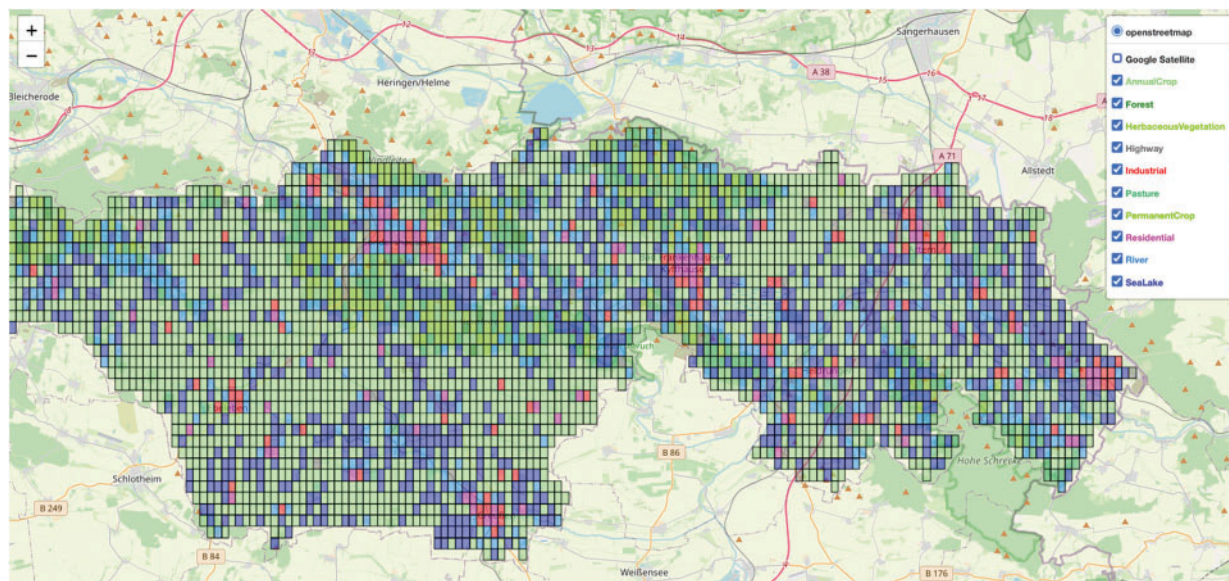
By fine-tuning the pre-trained ViT model, we achieved a state-of-the-art accuracy of 99.19% on the RGB bands of the EuroSAT dataset. This study reaffirms that transfer learning is an advantageous technique for LULC classification, as it allows for leveraging pre-existing knowledge for improved performance. We found that data augmentation techniques were helpful in mitigating overfitting and increasing dataset diversity, resulting in better model performance. Fine-tuning the training process with advanced techniques like learning rate optimization, regularization, and early stopping resulted in a significant performance boost while keeping resource demands under control.





**Figure 7:** Confusion matrix of non-augmented data with 20 epochs of ViT model

The study highlights the importance of careful model hyperparameter tuning for achieving optimal outcomes. The developed model can be used for mapping class distributions and extracting insights from geospatial imagery. We also identified several avenues for future research, including exploring the performance of other pre-trained models, incorporating additional data sources such as spectral bands or multi-temporal imagery, and investigating the applicability of these models to other LULC datasets. Focusing solely on class distribution mapping using RGB data has limited the investigation of the model’s full potential. Further research could explore its potential for other LULC-related tasks, such as change detection and land cover prediction.



**Figure 8:** Mapping of LULC of ‘Kreis Borken’ area using satellite images and ViT model

## 6 Conclusion

This research explored the application of transfer learning in LULC classification using the ViT model pre-trained on ImageNet-21K. By leveraging the transfer learning paradigm, we not only achieved outstanding accuracy (99.19%) on the EuroSAT RGB dataset but also established its role as a dependable and efficient method for LULC classification with deep learning models. We also observed that data augmentation improves the visual variability of training images, which in turn improves model performance. Moreover, we incorporated model enhancement techniques to optimize model training, improve performance, and reduce computational time. The most proficient model was subsequently used to map class distributions and offer insights into a specific region of geospatial imagery, which can assist in monitoring shifts in land usage and shaping policies for environmental conservation and urban development.

**Acknowledgement:** The authors would like to thank the editors and reviewers for their review and recommendations.

**Funding Statement:** The authors received no specific funding for this study.

**Author Contributions:** The authors confirm their contribution to the paper as follows: study conception and design: Suman Kunwar; experiment setup and analysis: Suman Kunwar; draft manuscript preparation: Suman Kunwar; manuscript verification and editing: Jannatul Ferdush. All authors reviewed the results and approved the final version of the manuscript.

**Availability of Data and Materials:** The authors confirm that the data supporting the findings of this study are available at <https://github.com/sumn2u/LULC-Mapping>.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

1. Mathieu E, Rodés-Guirao L. What are the sources for our world in data's population estimates? Our World in Data. 2022. Available from: <https://ourworldindata.org/population-sources> (accessed on 12/11/2023).
2. Levin K, Parsons S. 7 things to know about the IPCC's special report on climate change and land. 2019. Available from: <https://www.wri.org/insights/7-things-know-about-ipccs-special-report-climate-change-and-land> (accessed on 12/11/2023).
3. Huang M, Wang J, Li X. Land use change monitoring in nature reserves base on GF-1/GF-2. *2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)*. Shenyang, China: IEEE; 2021; p. 228–31.
4. Banhero S, de Abelleira D, Veron SR, Mosciaro MJ, Arévalos F, et al. Recent land use and land cover change dynamics in the gran Chaco Americano. *2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)*. Santiago, Chile: IEEE; 2020; p. 511–4.
5. Zhou W, Newsam S, Li C, Shao Z. PatternNet: a benchmark dataset for performance evaluation of remote sensing image retrieval. *ISPRS J Photogramm*. 2018;145(7553):197–209.
6. Milazzo VA. Role of spatial resolution and spectral content in change detection. *Proceedings-PECORA 9: Spatial Information Technologies for Remote Sensing Today and Tomorrow*. Sioux Falls, ND, USA: IEEE; 1984.
7. Hao H, Baireddy S, Bartusiak ER, Konz L, LaTourette K, et al. An attention-based system for damage assessment using satellite imagery. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. Brussels, Belgium: IEEE; 2021; p. 4396–9.
8. Sarath RN, Varghese JT, Bhatkar R. Effect of global climate change on land surface temperature over Dubai, United Arab Emirates with the aid of LANDSAT 8 satellite imagery. *8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*. Noida, India: IEEE; 2020; p. 1289–92.
9. Alex DAC, Antonio LOD. Review of methods for monitoring elements using satellite imagery. *17th International Conference on Engineering of Modern Electric Systems (EMES)*. Oradea, Romania: IEEE; 2023; p. 1–4.
10. Gladys Villegas R, van Coillie F, Ochoa D. Mapping and assessment of land use and land cover for different ecoregions of Ecuador using phenology-based classification. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. Brussels, Belgium: IEEE; 2021; p. 6492–5.
11. Sathyanarayanan D, Anudeep D, Keshav Das CA, Bhanadarkar S, Damotharan U, et al. A multiclass deep learning approach for LULC classification of multispectral satellite images. *2020 IEEE India Geoscience and Remote Sensing Symposium (InGARSS)*. Ahmedabad, Gujarat, India: IEEE; 2020; p. 102–5.
12. Waghela H, Patel S, Sudesan P, Raorane S, Borgalli R. Land use land cover classification using machine learning. *2022 International Conference on Automation, Computing and Renewable Systems (ICACRS)*. Pudukkottai, India: IEEE; 2022; p. 708–11.
13. Alem A, Kumar S. Deep learning methods for land cover and land use classification in remote sensing: a review. *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*. Noida, India: IEEE; 2020; p. 903–8.
14. Chen Y, Zheng L, Zheng J, Cheng J. Remote sensing image scene classification method based on semantic and spatial interactive information. *7th International Conference on Image, Vision and Computing (ICIVC)*. Xi'an, China: IEEE; 2022; p. 436–41.
15. Cheng G, Xie X, Han J, Guo L, Xia GS. Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities. *IEEE J Sel Top Appl*. 2020;13:3735–56.
16. Karra K, Kontgis C, Statman-Weil Z, Mazzariello JC, Mathis M, et al. Global land use/land cover with Sentinel 2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. Brussels, Belgium: IEEE; 2021; p. 4704–7.

17. Ojha SK, Challa K, Vemuri MK, Yarlagadda NSV, Phaneendra Kumar BLN. Land use prediction on satellite images using deep neural nets. *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*. Madurai, India: IEEE; 2019; p. 999–1003.
18. Taye MM. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Comput.* 2023;12(5):91.
19. Deng J, Dong W, Socher R, Li LJ, Li K, et al. ImageNet: a large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. Miami, Florida: IEEE; 2009; p. 248–55.
20. Helber P, Bischke B, Dengel A, Borth D. Eurosat: a novel dataset and deep learning benchmark for land use and land cover classification. *IEEE J Sel Top Appl.* 2019a;12(7):2217–26.
21. Temenos A, Temenos N, Kaselimi M, Doulamis A, Doulamis N. Interpretable deep learning framework for land use and land cover classification in remote sensing using shap. *IEEE Geosci Remote Sens Lett.* 2023;20:1–5.
22. Patel S, Ganatra N, Patel R. Multi-level feature extraction for automated land cover classification using deep cnn with long short-term memory network. *6th International Conference on Trends in Electronics and Informatics (ICOEI)*. Tirunelveli, India: IEEE; 2022; p. 1123–8.
23. Huo Y, Jin K, Cai J, Xiong H, Pang J. Vision transformer (Vit)-based applications in image classification. *2023 IEEE 9th International Conference on Big Data Security on Cloud (Big Data Security), IEEE International Conference on High Performance and Smart Computing (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS)*. New York, USA; 2023; p. 135–40.
24. Kumari M, Kaul A. Recent advances in the application of vision transformers to remote sensing image scene classification. *Remote Sens Lett.* 2023;14(7):722–32.
25. Jannat FE, Willis AR. Improving classification of remotely sensed images with the swin transformer. *SoutheastCon.* 2022;2022:611–8.
26. Zhai X, Kolesnikov A, Houlsby N, Beyer L. Scaling vision transformers. 2022. arXiv:2106.04560.
27. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, et al. An image is worth  $16 \times 16$  words: transformers for image recognition at scale. 2021. arXiv:2010.11929.
28. Fang B, Kou R, Pan L, Chen P. Category-sensitive domain adaptation for land cover mapping in aerial scenes. *Remote Sens.* 2019;11(22):2631.
29. Martini M, Mazzia V, Khaliq A, Chiaberge M. Domain-adversarial training of self-attentionbased networks for land cover classification using multi-temporal sentinel-2 satellite imagery. *Remote Sens.* 2021;13(13):2564.
30. Xue Z, Liu B, Yu A, Yu X, Zhang P, et al. Self-supervised feature representation and few-shot land cover classification of multimodal remote sensing images. *IEEE Trans Geosci Remote Sens.* 2022;60:1–18.
31. Capliez E, Ienco D, Gaetano R, Baghdadi N, Salah AH. Unsupervised domain adaptation methods for land cover mapping with optical satellite image time series. *IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium*. Kuala Lumpur, Malaysia: IEEE; 2022; p. 275–8.
32. Ruswurm M, Wang S, Korner M, Lobell D. Meta-learning for few-shot land cover classification. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Seattle, WA, USA: IEEE; 2020; p. 788–96.
33. Pradhan B, Al-Najjar HAH, Sameen MI, Tsang I, Alamri AM. Unseen land cover classification from high-resolution orthophotos using integration of zero-shot learning and convolutional neural networks. *Remote Sens.* 2020;12(10):1676.
34. Anil A, Variyar SVV, Sowmya V, Sukumar A, Krichen M, Vinayakumar R. Influence of spectral bands on satellite image classification using vision transformers. *TechRxiv.* 2022a. doi: 10.36227/techrxiv.20001764.v1.

35. Anil A, Variyar SVV, Sowmya V, Sukumar A, Krichen M, et al. Influence of spectral bands on satellite image classification using vision transformers. Aizawl, Mizoram, India: Springer Nature Singapore; 2022b. p. 243–51.
36. Helber P, Bischke B, Dengel A, Borth D. Introducing eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium*. Valencia, Spain: IEEE; 2018; p. 204–7.
37. Li J, Lin D, Wang Y, Xu G, Zhang Y, et al. Deep discriminative representation learning with attention map for scene classification. *Remote Sens*. 2020;12(9):1366.
38. Yassine H, Tout K, Jaber M. Improving LULC classification from satellite imagery using deep learning—EUROSAT dataset. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2021; p. 369–76.
39. Naushad R, Kaur T, Ghaderpour E. Deep transfer learning for land use and land cover classification: a comparative study. *Sens*. 2021;21(23):8083.
40. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data*. 2019;6(1):60.
41. Zhang T, Zhu T, Gao K, Zhou W, Yu PS. Balancing learning model privacy, fairness, and accuracy with early stopping criteria. *IEEE Trans Neural Netw Learn Syst*. 2023;34(9):5557–69.
42. Ede JM, Beanland R. Adaptive learning rate clipping stabilizes learning. *Mach Learn Sci Technol*. 2020;1(1):015011.
43. Dossa RFJ, Huang S, Ontañón S, Matsubara T. An empirical investigation of early stopping optimizations in proximal policy optimization. *IEEE Access*. 2021;9:117981–92.
44. He X, Li C, Zhang P, Yang J, Wang XE. Parameter-efficient model adaptation for vision transformers. 2023. arXiv:2203.16329.
45. Helber P, Bischke B, Dengel A, Borth D. Eurosatsat: a novel dataset and deep learning benchmark for land use and land cover classification. *IEEE J Sel Top Appl Earth Observ Remote Sens*. 2019b;12(7):2217–26.
46. Jia SJ, Wang P, Jia PY, Hu SP. Research on data augmentation for image classification based on convolution neural networks. *2017 Chinese Automation Congress (CAC)*. Jinan, China: IEEE; 2017; p. 4165–70.
47. Li J, Gao M, D’Agostino R. Evaluating classification accuracy for modern learning approaches. *Stat Med*. 2019;38(13):2477–503.
48. Zhang J, He T, Sra S, Jadbabaie A. Why gradient clipping accelerates training: a theoretical justification for adaptivity. 2020. arXiv:1905.11881.
49. Maji K, Gupta S. Evaluation of various loss functions and optimization techniques for mri brain tumor detection. *2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*. Ballari, Karnataka, India: IEEE; 2023; p. 1–6.
50. Mikołajczyk A, Grochowski M. Style transfer-based image synthesis as an efficient regularization technique in deep learning. *2019 24th International Conference on Methods and Models in Automation and Robotics (MMAR)*. Miedzyzdroje, Poland: IEEE; 2019; p. 42–7.
51. Awad M. Google earth engine (GEE) cloud computing based crop classification using radar, optical images and support vector machine algorithm (SVM). *2021 IEEE 3rd International Multidisciplinary Conference on Engineering Technology (IMCET)*. Beirut, Lebanon: IEEE; 2021; p. 71–6.